Lung Subtyping Signature for Predicting Survival in Lung Adenocarcinoma and Relationship to Other Gene Expression Prognostic Panels

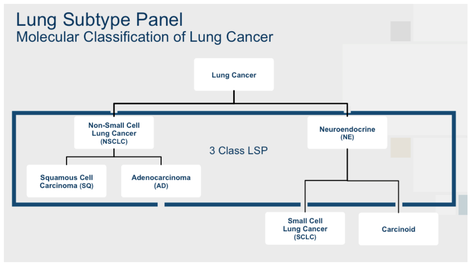
Mayhew, G[1], Hayes N[2], Perou C[2,3], Lai-Goldman, M[1], Faruki H.[1] (1)GeneCentric, Durham, NC (2)Lineberger Cancer Center, (3)Department of Genetics, University of North Carolina, Chapel Hill, NC

BACKGROUND

The Lung Subtype Panel (LSP), a previously validated 52-gene expression signature for classifying FFPE lung tumor samples into Adenocarcinoma (AD), Squamous Cell Carcinoma (SQ), and Neuroendocrine (NE) (NE comprising small cell carcinoma and carcinoid), was evaluated for its prognostic value in multiple publically available datasets. Prognosis by LSP was compared to several other prognostic gene expression panels, developed for Non Small Cell Lung Cancer (NSCLC) or Adenocarcinoma (AD). Survival predictions and prognostic strength, as compared to LSP, was examined.

METHODS

A previously published LSP 3-class nearest centroid predictor[1,2] was applied to stage I and II histology defined AD samples in the Director's Challenge[3] (Affy array, n=371), TCGA[4] (RNAseq, n=399), and Tomida et al.[5] (Agilent array, n=92) datasets.



Kaplan Meier plots and logrank tests were used to assess 5-year overall survival (OS) in two groups, those that were histologically and GE concordant (AD-AD) and those that were histologically and GE discordant (AD predicted SQ or NE (AD-NE/SQ)). Cox models were used to examine the LSP hazard ratio and to compare it with several other prognostic panels, Wilkerson et al.,[6] (506 genes) Wistuba et al.,[7] (31 genes) Kratz et al.,[8] (11 genes) and Zhu et al.,[9](15 genes). For Wistuba et al., genes were weighted equally. For Kratz et al., genes were weighted according to the coefficients in the publication. For Zhu et al., genes were weighted -1 or +1 according to the direction of effect on OS in the TCGA AD data set. For Wilkerson et al., the risk score was calculated as distance to the TRU (bronchioid) centroid. Gene mutation prevalence was examined for significantly associated mutations of lung AD[4] and SQ.[10]

RESULTS

LSP applied to 3 AD gene expression datasets. Each AD sample was assigned a Gene Expression-based (GE) subtype (AD-AD, AD-NE or AD-SQ) as shown in Table 1.

Table 1

Datasets	AD-AD	AD-NE	AD-SQ	Sum
Shedden et al.[3]	299 (0.81)	50 (0.13)	22 (0.06)	371
TCGA AD[4]	326 (0.82)	43 (0.11)	30 (0.08)	399
Tomida et al.[5]	72 (0.78)	12 (0.13)	8 (0.09)	92
Pooled set	697 (0.81)	105 (0.12)	60 (0.07)	862

Figure 1. Kaplan Meier plots and logrank tests evaluated in all histology and gene expression concordant samples (AD-AD) and discordant samples (AD-NE/SQ).

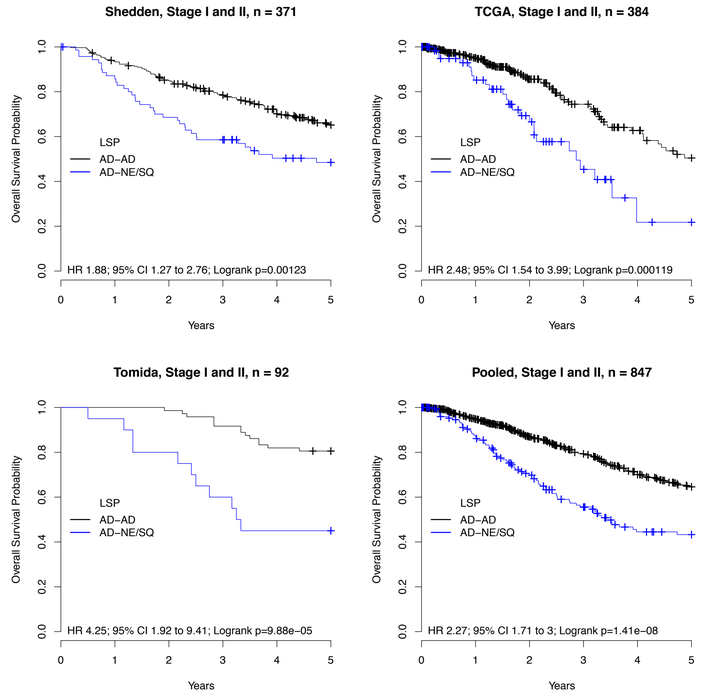


Table 2. Cox proportional hazard models of OS. Models in the hazard ratios table used binarized risk scores (at 0.67 quantile), calling one third of the samples high risk. Models in the p-values table left all risk scores continuous. All models adjusted for (T, N, Age).

Hazard Ratios	Covariates	AD-NE/SQ vs. AD-AD	Wilkerson[6]	Wistuba[7]	Kratz[8]	Zhu[9]	All
LSP AD-notAD vs AD-AD		2.16 (1.61,2.88)					1.6 (1.13,2.26)
Wilkerson et al. high vs low			1.73 (1.32,2.26)				1.1 (0.74,1.64)
Wistuba et al. high vs low				1.79 (1.38,2.33)			1.17 (0.78,1.74)
Kratz et al. high vs low					1.7 (1.3,2.22)		1.28 (0.93,1.77)
Zhu et al. high vs low						1.53 (1.17,2)	1.29 (0.97,1.7)
T Stage T2 vs T1	1.23 (0.92,1.64)	1.16 (0.87,1.55)	1.13 (0.84,1.51)	1.15 (0.86,1.54)	1.13 (0.84,1.52)	1.16 (0.87,1.55)	1.06 (0.79,1.42)
T Stage T3 or T4 vs T1	2.6 (1.53,4.4)	2.09 (1.22,3.56)	2.31 (1.36,3.94)	2.29 (1.35,3.88)	2.26 (1.32,3.85)	2.49 (1.47,4.21)	1.84 (1.07,3.16)
N Stage N1-N3 vs N0	2.43 (1.81,3.25)	2.46 (1.83,3.29)	2.37 (1.77,3.18)	2.45 (1.83,3.28)	2.37 (1.77,3.18)	2.34 (1.75,3.14)	2.39 (1.78,3.21)
age>65 vs age<=65	1.59 (1.22,2.07)	1.63 (1.25,2.12)	1.65 (1.27,2.15)	1.64 (1.26,2.13)	1.62 (1.25,2.11)	1.6 (1.23,2.08)	1.66 (1.27,2.16)
Male vs Female	1.27 (0.98,1.64)						
P-values							
LSP AD-notAD vs AD-AD		<0.0001					0.0189
Wilkerson et al. high vs low			<0.0001				0.1497
Wistuba et al. high vs low				<0.0001			0.7137
Kratz et al. high vs low					<0.0001		0.4523
Zhu et al. high vs low						<0.0001	0.0086
T Stage T2 vs T1	0.1635	0.3051	0.8568	0.6601	0.4764	0.4492	0.9914
T Stage T3 or T4 vs T1	0.0004	0.007	0.0097	0.007	0.0039	0.0023	0.0577
N Stage N1-N3 vs N0	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
age>65 vs age<=65	0.0006	0.0003	0.0001	0.0002	0.0004	0.0005	0.0002
Male vs Female	0.0759						

Figure 2. Significant differences in gene mutation prevalence in histology–gene expression concordant (AD-AD) as compared to discordant (AD-NE/SQ) samples using Fisher's exact test.

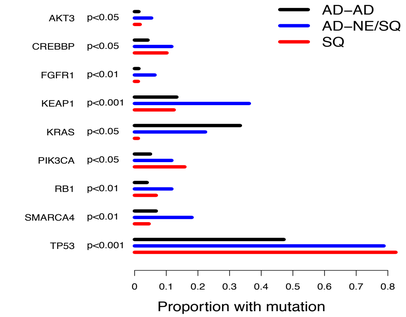
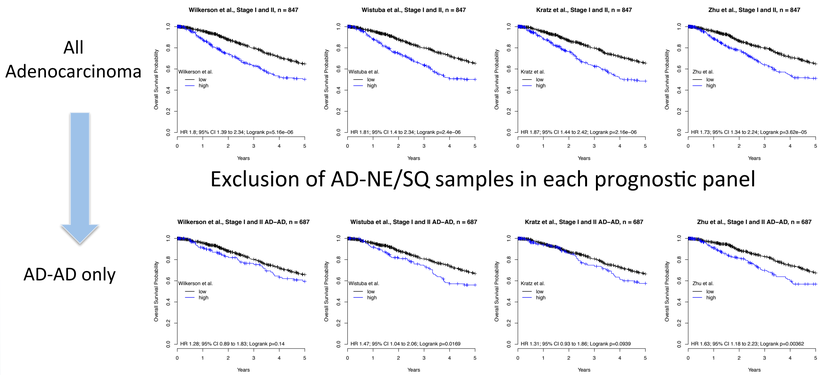


Figure 3. Reduction in lung adenocarcinoma prognostic strength following exclusion of histologically defined adenocarcinoma samples that are NE or SQ by LSP gene expression (AD-NE/SQ).



CONCLUSIONS

- ~20% histologically defined lung AD differ by gene expression subtype.
- Histology-GE discordant AD tumors demonstrate worse survival.
- Histology-GE discordant AD tumors are responsible for much of the prognostic risk in multiple prognostic gene signatures.
- Mutation frequencies in Histology-GE discordant samples differ significantly from concordant samples for 9/48 genes evaluated.
- Survival differences may be attributable to tumor biology and/or to variable response to standard AD management.

REFERENCES

1. Wilkerson MD, et al. J Molec Diag 2013; 15:485-497. PMID 23701907
2. Faruki H, et al. Archives Path & Lab Med. October 2015 PMID 26430809.
3. Shedden K. et al. Nat Med 2008; 14(8): 822-827. PMID 18641660
4. TCGA Lung AdenoC. Nature 2014; 511(7511): 543-550. PMID 25079552
5. Tomida S. et al. J Clin Oncol 2009; 27(17):2793-99. PMID 19414676
6. Wilkerson MD, et al. PLoS One 2012; 7(5): e36530. PMID 22590557
7. Wistuba II, et al. Clin Cancer Res 2013; 19(22):6261-6271. PMID 24048333
8. Kratz JR, et al. Lancet 2012; 379(9818):823-832. PMID 22285053
9. Zhu CQ, et al. J Clin Oncol 2010; 28(29):4417-4424. PMID 20823422
10. TCGA Lung SQCC. Nature 2012; 489(7417): 519-525. PMID 22960745

ABBREVIATIONS

AD =Adenocarcinoma
GE = Gene Expression
NE = Neuroendocrine (carcinoid or small cell carcinoma)
OS = Overall Survival
SQ = Squamous cell carcinoma
AD-AD =AD by both histology & GE
AD-NE/SQ = AD by histology & NE or SQ by GE

Abstract #ST49. Presented at the Association of Molecular Pathology Annual Meeting. Austin, TX. Nov 4-7, 2015